DOCUMENT RESUME

ED 464 125 TM 033 806

AUTHOR Fan, Xitao
TITLE Power of Latent Growth Modeling for Detecting Group Differences in Linear Growth Trajectory Parameters.
PUB DATE 2002-04-00
NOTE 36p.; Paper presented at the Annual Meeting of the American Educational Research Association (New Orleans, LA, April 1-5, 2002).
PUB TYPE Reports - Research (143) -- Speeches/Meeting Papers (150)
EDRS PRICE MF01/PC02 Plus Postage.
DESCRIPTORS Analysis of Variance; Groups; Sample Size; Simulation; *Statistical Bias; *Structural Equation Models
IDENTIFIERS *Latent Growth; Power (Statistics)

ABSTRACT

This simulation study focused on the power of detecting group differences in linear growth trajectory parameters within the framework of structural equation modeling (SEM) and compared this approach with the more traditional repeated measures analysis of variance (ANOVA) approach. Three broad conditions of group differences in linear growth trajectory were considered. SEM latent growth modeling consistently showed higher statistical power for detecting group differences in the linear growth slopes than repeated measures ANOVA. For small group differences in the growth trajectories, large sample size (e.g., N>500) is typically required for adequate statistical power. For medium or large group differences, moderate or small sample size is usually sufficient for adequate power. Some future research directions are discussed. (Contains 2 figures, 8 tables, and 16 references.) (Author/SLD)

Reproductions supplied by EDRS are the best that can be made
from the original document.

ED 464 125

Power of Latent Growth Modeling for Detecting Group Differences in Linear Growth Trajectory Parameters

Xitao Fan

University of Virginia

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)
☑ This document has been reproduced as received from the person or organization originating it.
☐ Minor changes have been made to improve reproduction quality.
• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

Running Head: Power in growth modeling

Send correspondence about this paper to:

Xitao Fan
Curry School of Education
University of Virginia
405 Emmet Street South, PO Box 400277
Charlottesville, VA 22904-4277

Phone: (434)243-8906
Fax: (434)924-1384
E-Mail: xfan@virginia.edu
WEB: http://www.people.virginia.edu/~xf8d

PERMISSION TO REPRODUCE AND DISSEMINATE THIS MATERIAL HAS BEEN GRANTED BY
X. Fan
TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

TM033806

Paper presented at the 2002 annual meeting of the American Educational Research Association, April 3 (*Session # 26.24*), New Orleans, LA

BEST COPY AVAILABLE

Abstract

This simulation study focused on the power for detecting group differences in liner growth trajectory parameters within the framework of structural equation modeling (SEM), and compared this approach with more traditional repeated measures ANOVA approach. Three broad conditions of group differences in linear growth trajectories were considered. SEM latent growth modeling consistently showed higher statistical power for detecting group differences in the linear growth slopes than repeated measures ANOVA. For small group differences in the growth trajectories, large sample size (e.g., $\underline{N} > 500$) is typically required for adequate statistical power. For medium or large group differences, moderate or small sample size is usually sufficient for adequate power. Some future research directions were discussed.

KEYWORDS: latent growth modeling, power, longitudinal data analysis, Monte Carlo studies

Developmental trends in human behaviors have long been an area of research interest in social science disciplines (e.g., education, psychology, sociology). To accommodate such substantive research interest, different analytic techniques have been developed to help substantive researchers to understand developmental behavioral change. For example, repeated measures analysis of variance is often used for longitudinal data with multiple measurements. Auto-regressive models, either in the form of regression analysis, or within the framework of structural equation modeling, has also been popular for longitudinal analysis. More recently, individual change has been studied within the general framework of growth modeling

Growth modeling for studying individual change may be implemented either within the framework of hierarchical linear modeling (e.g., Bryk & Raudenbush, 1992), or within the framework of structural equation modeling (e.g., Willet & Sayer, 1996). Under structural equation modeling (SEM), modeling growth for individual change may take one of several slightly different names, such as latent growth modeling, latent growth curve analysis, or latent trajectory models. Further, SEM approach for modeling individual growth has probably become the method of choice for understanding developmental change. Given that SEM is often considered as a very general analytic model that subsumes many other analytical techniques (e.g., Fan, 1996), it is not surprising that SEM plays an important role in longitudinal data analysis for developmental trend. Of the many issues in the application of latent growth modeling, the power of this analysis for detecting potential group differences in latent growth trajectory parameters is an area where our understanding appears to be limited.

The purpose of this paper is to present the results of an empirical study in which the power of latent growth modeling (LGM) for detecting group differences in the growth trajectory parameters (i.e., intercept and slope of the growth trajectory) are empirically assessed. When

applicable, the results from LGM are compared with the traditional repeated measures ANOVA in terms of both Type I error rate and the power.

Perspectives

Traditional Approaches for Time-Ordered Measurement Series

Assessing behavioral developmental growth requires multiple measurements from the same individuals over a finite time span. Traditionally, auto-regressive cross-lagged model, either in the form of a series of regression analysis, or in the form of structural equation modeling, has often been the analytic technique for data of repeated measurements. Auto-regressive model focuses on the relations between adjacent measurements, but not on the overall growth trajectory as represented by all the repeated measurements. In addition, auto-regressive cross-lagged model generally does not pay attention to the mean structure of the growth trajectory.

Repeated measure ANOVA is another popular technique for analyzing data of multiple measurements. For repeated measurements from a single group, repeated measure ANOVA focuses on the measurement change over time (within-subject factor). For repeated measurements from multiple groups, repeated measure ANOVA focuses on three aspects of the data structure: the difference of the overall group means across the repeated measurements (between-subject factor), the measurement change over time (within-subject factor), and the interaction of the between-subject factor (group membership) and the within-subject factor (time: measurement change over time). If it can be assumed that there exists a linear developmental trend in the data structure over the repeated measurements, the within-subject factor (measurement change over time) is conceptually comparable to detecting if there is growth over repeated measurements. On the other hand, the interaction of the between-subject factor (group membership) and within-subject factor (Time) is conceptually comparable to testing the

hypothesis that the growth rate over repeated measurements is equal for multiple groups.

Although repeated measure ANOVA models both the change over time (growth), and any potential group difference in the change over time (different growth patterns for groups), it does not provide sufficient information about the growth trajectories (cf. Bryk & Raudenbush, 1992; Ware, 1985). More specifically, repeated measure ANOVA does not model the starting point (intercept) of a growth trajectory, and nor does it model the rate of change in the process.

Modeling Individual Growth

Modeling growth within the SEM framework is a more recent approach for studying developmental trends based on a series of measurements. Because SEM latent growth modeling (LGM) approach offers more flexibility in testing different research hypotheses about the developmental trend, many researchers have argued in favor of its superiority over the traditional analytic approaches for modeling developmental trends (e.g., Curran, 2000; Duncan, Duncan, Strycker, Li, & Alpert, 1999; McArdle & Bell, 2000).

Unconditional growth model. Assuming a series of repeated measurements Y_{it} (minimum three repeated measurements, Y_{i1}, Y_{i2}, Y_{i3}; i represents an individual, and t represents the time-ordered measurements of Y), the growth model for describing an individual's growth as represented by this series of repeated measurements is called the level 1 or within-person model:

$$Y_{it} = \alpha_i + \beta_i \lambda_t + \varepsilon_i \qquad (1)$$

where α_i represents the intercept of an individual's growth trajectory (i.e., the initial status measured at Time 1), β_i represents the slope of an individual's growth trajectory (i.e., the unit change in Y_i between two consecutive measurements), λ_t represents consecutive time points at which the measurement is taken, and ε_i represents the modeling residual for an individual.

Because the intercept (α_i) and the slope (β_i) are random variables, these individual model parameters can be represented by the group mean intercept (μ_α) and group mean slope (μ_β) plus individual variation ($\zeta_{\alpha i}$, $\zeta_{\beta i}$) in the following level 2 or between-person model:

$$\begin{aligned} \alpha_i &= \mu_\alpha + \zeta_{\alpha i} \\ \beta_i &= \mu_\beta + \zeta_{\beta i} \end{aligned} \tag{2}$$

The Level 2 model assumes that an individual growth trajectory parameter is the function of the group mean growth parameter plus individual variation, and no other predictors are involved in the model to account for the variation of individual intercepts and slopes. The Level 2 model presented above is often called the unconditional model (e.g., Curran, 2000). The unconditional latent growth model (for three repeated measurements) is represented graphically as an SEM model in Model A in Figure 1.

Insert Figure 1 about here

Conditional growth model. In situations where research calls for testing for a predictor (X_1) that may explain the variation of the individual growth trajectory parameters (i.e., α_i and β_i), a conditional Level 2 (between-person) model can be constructed (Curran, 2000):

$$\begin{aligned} \alpha_i &= \mu_\alpha + \gamma_1 X_{1i} + \zeta_{\alpha i} \\ \beta_i &= \mu_\beta + \gamma_2 X_{1i} + \zeta_{\beta i} \end{aligned} \tag{3}$$

where, X_1 is the predictor that is hypothesized to affect the individual growth trajectory parameters (α_i and β_i), an γ_1 and γ_2 are the path coefficients representing the systematic effects

of X_1 on the variation of an individual growth trajectory parameters of intercept (α_i) and slope (β_i) respectively. For example, we may ask if X_1 affects the initial status (i.e., the intercept) of the growth trajectory, or if it accelerates the growth (i.e., steeper slope for the trajectory). The conditional latent growth model (for three repeated measurements) is graphically represented as an SEM model as Model B in Figure 1.

Power for Detecting Group Differences in Latent Growth Model

In SEM, different approaches may be used for detecting potential group differences in the growth trajectory parameters of a latent growth model. A very general approach is to conduct multi-sample analysis in structural equation modeling. In this approach, the same growth model is implemented for separate group sample data, and equality constraints representing different research hypotheses, e.g., equality of intercepts and/or equality of growth slopes, can be imposed. Hierarchical χ^2 test for the nested models (i.e., multi-sample analyses with and without the constraints of interest) can be used for assessing the potential group differences on the latent growth trajectory parameters.

The research work by Muthén and Curran (1997) probably represents the pioneering work specifically for power estimation in detecting group difference in latent growth model trajectory, and their work was based on the earlier work in the area of power of SEM in detecting misspecified models (Satorra & Saris, 1985; Saris & Satorra, 1993; also see Kaplan, 1995, for a concise review and discussion of relevant issues). In Muthén and Curran's (1997) work, they focused on the experimental design of two groups (control vs. treatment groups), and detecting the difference in the growth slope parameter only (i.e., growth rate difference) was their focus. For this purpose, they implemented the models of the same intercept (representing randomized

design with equal starting point) and the same slope for the two groups, but for the treatment group, they added an additional growth parameter to model the accelerated growth of the treatment group compared with the control group. Duncan et al. (1999, Chapter 10) provide the procedural details for this approach of power estimation in latent growth curve modeling.

In most educational and psychological research situations, either because of practical constraints (e.g., practically impossible for randomized design), or because of the nature of research (research focus on naturally occurring and intact groups, such as male vs. female), quasi-experimental or non-experimental design is called for (Cook & Campbell, 1979; Shadish, Cook, & Campbell, 2001). In these situations, The approach of using an added growth factor as illustrated in Muthén and Curran (1997) and Duncan et al. (1999) is not directly applicable, because the groups may potentially differ in the initial status (the intercept of the growth trajectory), in the growth rate (the slope of the growth trajectory), or both.

In the situation where comparison of two groups is being made, the group membership can be readily dummy coded, and the dummy coded group membership can be used as the predictor ($\underline{X_1}$) for the growth trajectory intercept and slope, i.e., $\underline{\alpha_i}$, or $\underline{\beta_i}$, or both, as described graphically in Model B of Figure 1. Consequently, potential groups differences in the growth trajectory parameters can be operationalized as the coefficients of $\underline{\gamma_1}$ and $\underline{\gamma_2}$ in Equation 3, and such potential group differences can be tested through the statistical assessment of $\underline{\gamma_1}$ and $\underline{\gamma_2}$, for intercept and slope differences respectively.

While the approach by Muthén and Curran (1997) relies on the global statistical test for misspecified SEM model, the approach of using dummy coded group membership described here relies on the local statistical tests for two specific parameter estimates ($\underline{\gamma_1}$ and $\underline{\gamma_2}$ in Equation 3).

To operationalize group differences as local parameters may have some advantages over the approach of relying on the global test for misspecified models for detecting potential group differences. One obvious advantage is that this approach may offer more flexibility for testing different hypotheses. For example, we may be interested in any potential group differences (e.g., ethnic group differences) in growth trajectory parameters after controlling for one or more covariates (e.g., SES). In this case, the covariate to be controlled for can be operationalized in Model B of Figure 1 as another predictor (e.g., X_2, parallel to the dummy coded group membership X_1) for growth trajectory parameters. The potential group differences in the intercept and slope of the growth trajectory are readily assessed by testing for the path coefficients of γ_{11} and γ_{21} in the following Equation 4.

$$\begin{aligned} \alpha_i &= \mu_\alpha + \gamma_{11} X_{1i} + \gamma_{22} X_{2i} + \zeta_{\alpha i} \\ \beta_i &= \mu_\beta + \gamma_{21} X_{1i} + \gamma_{22} X_{2i} + \zeta_{\beta i} \end{aligned} \qquad (4)$$

This paper examined the issue of statistical power of LGM in SEM in detecting group differences in linear growth trajectory parameters within the context of non-experimental design, and assumed that the group differences may potentially occur in both growth trajectory parameters (intercept, slope). The study implemented a Monte Carlo simulation design for assessing the power of LGM for detecting group differences in latent growth trajectory parameters, and the Type I error rate when group differences did not exist. When applicable, the latent growth modeling approach was compared with results of the repeated measures ANOVA.

Methods

Simulation Design

In this study, two groups' growth trajectories based on five equally spaced repeated

measures were simulated. The five repeated measures formed a linear growth trajectory, and curvilinear trajectory conditions were not considered in this study. Further, data non-normality was not considered. The five repeated measures had a simplex covariance structure pattern, with temporally closer measurements being more highly correlated than temporally farther measurements. Three patterns of group differences in latent growth trajectory parameters were simulated: (1) group difference in intercept only, (2) group difference in slope only, and (3) group differences in both growth trajectory intercept and slope.

For the first pattern that group difference only exists in the intercept of the growth trajectory, four conditions were simulated: group differences equal to zero, small, medium, and large effect sizes as measured by the standardized mean differences between the two groups ($\underline{d}$ = 0.0, 0.2, 0.5, and .8 respectively; see Model I in Figure 2). For the second pattern that group difference exist only in the slope of the growth trajectory, three conditions are simulated: two groups started with equal growth trajectory intercept (1st measurement), but end (5th measurement) with differences that were equal to small ($\underline{d}$ = 0.2), medium ($\underline{d}$ = 0.5) and large ($\underline{d}$ = .8) effect sizes (Model II in Figure 2).

Insert Figure 2 about here

For the third pattern that group differences exist both in the intercept and in the slope of the growth trajectory, three situations were simulated. In the first situation, two groups started (1st measurement) with a small intercept difference ($\underline{d}$ = 0.2), but ended (5th measurement) with differences that were equal to medium ($\underline{d}$ = 0.5) and large ($\underline{d}$ = .8) effect sizes respectively. This was a situation where the difference between the groups increased with repeated measurements, and the group with higher intercept had the accelerated growth slope compared with the other

group with lower intercept. This situation is graphically depicted as Model III in Figure 2.

Model IV in Figure 2 graphically depicts the second situation under the pattern that group differences exist both in the intercept and the slope of the growth trajectory. In this situation, the two groups started with a intercept difference that were equal to a medium (d = .5) or large (d = .8) effect size. The group with lower intercept had slightly faster growth, and the two groups ended with a difference of a small effect size (d = .2).

The third situation under the pattern that group differences existed both in the intercept and the slope of the growth trajectory is shown by Model V in Figure 2. In this situation, there is a disordinal group × growth interaction pattern. The two groups started with a small difference (d = .2) in the intercepts. The group with the lower intercept had faster growth. In the end, the group with the lower intercept surpassed the other group, and ended with a difference equivalent to a small (d = .2) or medium (d = .5) effect size. The details for the simulated data conditions of the five models described above are presented in Table 1.

Insert Table 1 about here

Eleven sample size conditions (two equal groups combined) of 50, 100, 200, 300, 400, 500, 600, 700, 800, 900, and 1,000 were implemented for every growth trajectory conditions discussed above, and 2,000 replications were implemented within each cell condition. The design entailed a total of 286,000 $\{[(4 + 3 + 2 + 2 + 2) \times 11] \times 2000\}$ replications.

Data Source and Analyses

Data were simulated using a combination of SAS macro, SAS BASE, SAS PROC IML (Interactive Matrix Language). Latent growth models were implemented through SAS PROC CALIS (Covariance Analysis of LInear Structures), and repeated measures ANOVA were

implemented through SAS PROC GLM (General Linear Model procedure). In general, repeated measures ANOVA does not provide the statistical mechanism for testing for group difference in the intercept of the growth trajectory. For group difference in the slope of the growth trajectory, the hypothesis that two groups have unequal growth trajectory slopes is statistically tested by the interaction term of the between-subject factor (i.e., group membership) and the within-subject factor (i.e., repeated measures). Repeated measures ANOVA provides both multivariate and univariate tests for this interaction term. In this study, both the multivariate test (based on the Wilk's Lambda statistic) and the univariate test for this interaction term were used in the comparison with the results from SEM latent growth modeling tests for the group effect (γ_2 in Equation 3) on the slope of growth trajectory.

Results and Discussions

Preliminary Results

When using SEM approach for modeling linear growth as represented in the model in Figure 1, there can be the problem of non-convergence. Non-convergence was empirically examined for the data conditions and the models implemented in this study. Table 1 presents the percentages of non-converging samples for different sample size conditions under each of the five models in Figure 2. In general, non-convergence does not appear to be a problem for the models implemented and the data conditions studied. For Models I to IV, only about less than one percent of the samples failed to converge when sample size was relatively small (e.g., N = 50, 100). Once sample size reached 200, almost all samples converged, and converged quickly, as indicated by the number of iterations (not presented in Table 2) before reaching convergence.

The Model V (disordinal interaction pattern between group membership and growth slope), however, stands out for having a non-negligible percentages of samples that failed to

converge, even when the sample size became large for the relatively simple linear growth model implemented (e.g., N = 300, 400). A closer examination revealed that a much larger proportion of non-convergent samples occurred for the condition of d = 0.50 at the fifth measurement. For example, for N = 100, for $\underline{d}$ = 0.20 and $\underline{d}$ = 0.50, the percentages of non-converging samples were 2.55% vs. 10.55% respectively. For N = 300, for $\underline{d}$ = 0.20 and $\underline{d}$ = 0.50, the percentages of non-converging samples were 0.35% vs. 4.55% respectively. These findings here appear to suggest that modeling this type of disordinal interaction pattern between the between-subject factor (group membership) and the within-subject factor (Time) in SEM may be potentially more difficult.

Type I Error When Group Difference Does not Exist

For Model I, when the intercept difference is zero, the two groups have no difference in either growth trajectory intercept or trajectory slope. Under these conditions, the Type I error rates of different analytic approaches for falsely declaring group difference were examined. As discussed previously, repeated measures ANOVA generally does not offer a statistical mechanism for testing potential group difference in the intercept of the growth trajectory. However, when the groups have parallel linear growth, i.e., when the two groups have the same growth slopes as in Model I, the univariate test for the between-subject factor (group) is equivalent to a test for the group intercept difference. Under the condition of parallel slopes, testing for group difference across the repeated measurements is the same as testing for the difference at the initial measurement (intercept). So under the condition of parallel growth slopes in Model I, repeated measures ANOVA does provide a univariate test for group intercept difference.

Figure 3 presents a graphic description for the Type I error rates of LGM and repeated measures analyses in falsely declaring group difference in the intercept and slope of the growth

trajectories when no difference exists. It is observed that, both for the intercept (SEM, univariate repeated measures) and for the slope (SEM, univariate repeated measures, and multivariate repeated measures), the empirical Type I error rates of different approaches are almost all within the range between 0.04 and 0.06, very close to the nominal Type I error rate ($\alpha = 0.05$). There is no indication that one approach is systematically inferior in controlling for Type I error.

Insert Figure 3 about here

Power for Detecting Intercept Difference in Model I

Figure 4 presents the power curves for detecting group differences in growth trajectory intercept when the trajectory slopes are the same for the two groups. For repeated measures ANOVA, as discussed above, under the condition of parallel growth slopes, the univariate test is applicable for testing group difference in growth trajectory intercept. Figure 4 shows that, univariate repeated measures analysis has more power than SEM modeling approach for detecting group differences in the growth trajectory intercept, under the condition that the growth trajectory slopes are equal for the groups. This pattern is observed for the three intercept difference conditions ($\underline{d}$ = 0.20, 0.50, 0.80). For example, when the group difference in the intercept is small ($\underline{d} = 0.20$), it will take SEM modeling approach a sample about 900 to reach statistical power of 0.80 (often considered as representing adequate power, see Cohen, 1988), while it takes repeated measures ANOVA a sample of 700 to reach the same power level.

Insert Figure 4 about here

Although for small group difference in intercept ($\underline{d} = 0.20$), it may take a large sample for either approach (SEM and univariate repeated measures) to detect such group differences, these

two approaches appear to have adequate power when group difference became larger (d = 0.50, 0.80), and it only takes a small to moderate sample size (N = 100 for d = 0.50, and N = 50 for d = 0.80) for the two analytic approaches to reach adequate power levels of 0.80.

Power for Detecting Slope Differences in Model II

In Model II, the two groups have equal intercepts, but have different growth rates, and consequently, the two group end with a difference of a small, medium, or large effect size. Figure 5 presents the power curves of three approaches (SEM modeling, multivariate repeated measures ANOVA test, univariate repeated measures ANOVA test) for detecting the group difference in the growth slopes.

Insert Figure 5 about here

Contrary to the observation in Figure 4 where the (univariate) repeated measures analysis showed more power for detecting group difference in the intercept of the growth trajectory, in Figure 5, LGM was more powerful than both the univariate repeated measures and multivariate repeated measures analyses for detecting group differences in the slope of the growth trajectory. This pattern is consistent for the three conditions of group slope differences, except when all three tests reached maximum power. It is also noted that the univariate repeated measures ANOVA test is consistently more powerful than the multivariate repeated measures ANOVA for detecting the group difference in growth slope, i.e., for detecting the interaction of the between-subject factor (group membership) and the within-subject factor (Time).

The difference in power between the univariate and multivariate repeated measures ANOVA tests is not surprising, however. Given the simulated data conditions, the population variances for the groups are equal, and so are the population covariances of the groups. In other

words, the data satisfy a much stronger condition of compound symmetry, a more stringent condition than that of sphericity that is typically assumed for repeated measures analysis (Stevens, 2002). As discussed in Stevens (2002), "if sphericity holds, then the univariate approach is more powerful. When sphericity is violated, however, then the situation is much more complex." (p. 509). It would be interesting to examine the performance of the univariate and multivariate repeated measures ANOVA in detecting the group difference in growth slope when the sphericity assumption is violated.

All three tests (LGM, univariate repeated measures, multivariate repeated measures) did not do well for detecting the small group difference in growth slopes as operationalized in Model II. For sample sizes of smaller than 400, all three tests have less than 50/50 chance for detecting such group difference. It will take a sample of about 700-800 for LGM approach to reach the power of 0.80, while it will take multivariate repeated measures ANOVA over 1000 (not simulated) to reach that power level. The power for detecting medium group difference in growth slopes as defined in Model II is generally adequate for the three tests, with SEM reaching the power level close to 0.80 with a sample a little over 100, and the multivariate repeated measures ANOVA reaching that level with a sample of 200. For the condition of large group difference in growth slopes as defined in Model II, all three tests quickly reached the maximum power level when the sample size reaches 100.

Power for Detecting Group Differences in Model III

Model III as represented in Figure 2 has one condition of small group difference in the growth intercept, and two conditions of group differences in growth slopes. Because the growth slope differences exist, repeated measures ANOVA could only be used for testing for group difference in growth slopes, but not for testing the group difference in growth intercepts. Figure

6 presents the power curves of the three tests for the two conditions of group difference in growth slopes (LGM, univariate repeated measures, multivariate repeated measures), and the power curve of LGM for the one condition of group difference in growth intercepts.

Insert Figure 6 about here

First, for LGM to reliably detect small group difference (as operationally defined in Model III) in the intercept of the growth trajectory, it will take a large sample size (approximately 600 to 700) to reach the power level of 0.80. This is similar to, although slightly better than, what is presented in Figure 4.

For detecting group difference in growth rates (growth trajectory slopes), the three tests (SEM, univariate and multivariate repeated measures ANOVA) showed the same pattern of performance differences as in Figure 5, with the multivariate repeated measures analysis has the least power, and LGM has the most power. The difference can be practically important. For example, for the first condition of slope difference (5th measurement end with a difference of $\underline{d}$ = 0.50), it will take a sample of 300 for SEM modeling to reach the power level of about 0.75, while it will take a sample of 500 for multivariate repeated measures analysis to reach the same power level. In terms of efficiency, LGM will be 166% [(500/300)*100%] more efficient than the multivariate repeated measures anlaysis for detecting such group difference in growth trajectory slopes. In research practice, this difference may mean a lot of savings in resources.

For the second condition of larger group difference in growth trajectory slopes, the three tests reached adequate power level (e.g., 0.80) with moderate sample size (e.g., $\underline{N}$ = 100). The same difference pattern in power exists among the three tests, except when all the three tests reached maximum power level (e.g., when $\underline{N}$ = 300).

Power for Detecting Group Differences in Model IV

In Model IV, the two groups started with a medium or large difference (intercept), and the lower performing group accelerated faster, and the two groups ended with a small difference. So in this model, there were two conditions of group intercept difference, and two conditions of group slope difference.

Because there was interaction of the between-subject factor (group) and the within-subject factor (Time), repeated measures ANOVA could not be used for testing the group intercept differences, and only LGM could serve this purpose. For both conditions of intercept differences ($\underline{d} = 0.5$, $\underline{d} = 0.8$), the SEM latent growth model showed adequate power for detecting such group differences (not presented graphically). More specifically, for the condition of $\underline{d} = 0.5$, SEM latent growth model reached power of 0.45 when sample size was 50, and power of 0.75 for sample size of 100, and to 0.95 for sample size of 200. For the condition of intercept difference of $\underline{d} = 0.80$, LGM reached the power level of 0.85 for sample size of 50, and reached maximum power when sample size was 100 and above.

For the two conditions of group slope difference, Figure 7 presents the power curves of the three tests for detecting the group differences in the slope of the growth trajectory. The pattern of difference among the three tests is similar to those in Figure 5 and Figure 6: LGM has the highest statistical power, which is closely followed by the univariate repeated measures analysis, with the multivariate repeated measures ANOVA test having the least statistical power. For the first condition of group slope difference (two groups started with $\underline{d} = 0.5$, ended with $\underline{d} = 0.2$), large sample size was needed for having adequate power. For example, for LGM, sample size about 300 to 400 was needed for reaching the power level of 0.8; for multivariate repeated measures, sample size about 500 to 600 was needed for the same power level. For the second

condition of group slope difference (two groups started with $\underline{d} = 0.8$, ended with $\underline{d} = 0.2$), all three tests quickly reached maximum power when sample size approaches 200.

Insert Figure 7 about here

Power for Detecting Group Differences in Model V

Model V simulated a disordinal interaction pattern of the between-subject factor (group membership) and the within-subject factor (Time). There was one condition of group difference in the intercept of the growth trajectory (started with $\underline{d} = 0.2$), and two conditions of group differences in the slope of the growth trajectory (ended with $\underline{d} = 0.2$ and $\underline{d} = 0.5$ in reversed direction). As discussed previously (see Preliminary Results section), this disordinal interaction pattern appeared to have caused some technical difficulties for LGM in the sense that, even when the sample size was relatively large, there was still a non-negligible proportion of non-converging samples for LGM (see Table 2 and related discussion). Figure 8 presents the power curves for testing the one condition of group difference in the intercept of the growth trajectory (SEM only; repeated measures approach was not applicable), and for the two conditions of group differences in the slope of the growth trajectory (three tests: SEM, univariate repeated measures, multivariate repeated measures). For LGM results, only the converged samples were usable for our discussion of power, and for the construction of the power curves.

Insert Figure 8 about here

For detecting group difference in the intercept of the growth trajectory, the power curve of SEM latent growth model is basically the same as the one in Figure 6, and it may require a large sample (e.g., $\underline{N} > 600$) for reliably detecting such a group difference ($\underline{d} = 0.2$) in the

intercept of the growth trajectories. For detecting group differences in the slope of the trajectories, because the groups ended with a difference in the reversed direction compared to their difference at the beginning, one group's slope had to be quite different from that of the other, and consequently, the group differences in the tarjectory slope in Model V were relatively large.

The pattern of differences in power among the three tests (SEM, univariate repeated measures, multivariate repeated measures) is essentially the same as before in Figures 5, 6, and 7, with LGM being the most powerful, and the multivariate repeated measures analysis being the least powerful. For the condition of the smaller slope difference in Model V, relatively large sample size is needed for adequate power (e.g., SEM: N = 200 for power of 0.80; multivariate repeated measures: N = 300 for power of 0.80). For the condition of the larger slope difference in Model V, the tests quickly reached maximum power when sample is moderate (e.g., N > 100), especially for LGM and univariate repeated measures ANOVA test.

Summary and Conclusions

This study focused on the power of detecting group differences in growth trajectory parameters within the framework of structural equation modeling. This study simulated three broad conditions of group differences in linear growth trajectories: group differences (a) only in the intercepts of the growth trajectories, (b) only in the slopes of the growth trajectories, and (c) both in the intercepts and slopes of the growth trajectories. For Condition (c), three different models representing some variation of this condition were simulated. Different degrees of group differences in the intercepts and the slopes were examined. Within each specific condition, 2,000 replications were simulated to ensure reasonably accurate estimation for power. A total of 286,000 replication samples were drawn from the statistically defined populations. The simulation design allowed systematic assessment of power of LGM in modeling growth

trajectories, and some comparisons of LGM power with more traditional repeated measures analysis approach. The findings provided some useful information for our better understanding of the relevant issues in LGM. The major findings from this study are as follows:

1. For both the intercept and slope of the growth trajectories, the empirical Type I error rates are very close to the nominal Type I error rates, and no test (SEM, univariate repeated measures, multivariate repeated measures) appeared to have shown any observable deficiencies in controlling for Type I error.

2. Under the condition of no group differences in the slopes of the growth trajectories, univariate repeated measures analysis showed more power for detecting the group differences in the intercepts of the growth trajectories than LGM.

3. LGM consistently showed more statistical power for detecting group differences in the slopes of the growth trajectories than repeated measures analysese. Univariate repeated measures showed more statistical power for detecting such group differences in the growth slopes than the multivariate repeated measures analysis, consistent with previous research when the data of the repeated measures satisfy the sphericity condition.

4. In general, for reliably detecting small group differences (as operationally defined in this study; see Table 1 and Figure 2) in either the intercepts and/or the slopes of the growth trajectories, large sample size is typically required. For example, $\underline{N} > 500$ (two groups combined) is typically required for power level of 0.70 - .80. For reliably detecting medium group differences in the intercepts and/or the slopes of the growth trajectories, moderate sample size, e.g., $\underline{N} = 100$ to 200, is typically sufficient for reaching adequate power level (0.70 to 0.80).

5. For detecting group differences in the slopes of the growth trajectories, the relative efficiency of LGM over the multivariate repeated measures analysis is typically within the range of

150% to 200%. That is, with SEM latent growth modeling approach, the sample size can be 2/3 to 1/2 of that for the multivariate repeated measures analysis for comparable statistical power.

6. For the data condition of Model V, i.e., the disordinal interaction pattern of the between-subject factor (Group) and within-subject factor (Time), there appeared to be some potential difficulty for LGM approach, because a non-negligible proportion of samples failed to converge even when sample size was relatively large. This suggests that this type of differences in the groups' growth trajectories might be more difficult for LGM.

Future research may consider different data conditions, especially data conditions that are NOT as "clean" as simulated in this study. For example, research in developmental trends shows that later measurements typically have larger variances than early measurements (for examples, see data in Fan, 2001; McArdle & Bell, 2000). In other words, individual differences typically become more pronounced as they develop. This type of more realistic data structures may be considered in simulation work in growth modeling. Future research in this area may also consider the comparison of SEM latent growth modeling approach with growth modeling under the framework of hierarchical linear modeling (HLM; Bryk & Raudenbush, 1992). It appears that no systematic research has been conducted for such and similar comparisons in terms of power and Type I error.

References

Bryk, A. S., & Raudenbush, S. W. (1992). Hierarchical linear models: Applications and data analysis methods. Newbury Park, CA: SAGE Publications.

Cohen, J. (1988). Statistical power analysis for the behavioral sciences (2nd ed.). Hillsdale, New Jersey: Lawrence Erlbaum Associates, Publishers.

Cook, T. D., & Campbell, D. T. (1979). Quasi-experimentation: Design and analysis issues for field studies. Chicago: Rand McNally.

Curran, P. J. (2000). In J. S. Rose, L. Chassin, C. C. Presson, & S. J. Sherman (Eds.), Multivariate applications in substance use research: New methods for new questions (pp. 1-42). Mahwah, New Jersey: LEA.

Duncan, T. E., Duncan, S. C., Strycker, L. A., Li, F., & Alpert, A. (1999). An introduction to latent variable growth curve modeling. Mahwah, New Jersey: LEA.

Fan, X. (1996). Structural equation modeling and canonical correlation analysis: What do they have in common? Structural Equation Modeling: A Multidisciplinary Journal, 4, 64-78.

Fan, X. (2001). The effect of parental involvement on high school students' academic achievement: A growth modeling analysis. Journal of Experimental Education, 70, 27-61.

Kaplan, D. (1995). Statistical power in structural equation modeling. In R. H. Hoyle (Ed.), Structural equation modeling: Concepts, issues, and applications (pp. 100-117). Thousand Oaks, CA: SAGE Publications.

McArdle, J. J., & Bell, R. Q. (2000). An introduction to latent growth models for developmental data analysis. In T. D. Little, K. U. Schnabel, and J. Baumert (Eds.), Modeling longitudinal and multilevel data (pp. 69-108). Mahwah, New Jersey: LEA.

Muthén, B., & Curran, P. (1997). General growth modeling of individual differences in

experimental designs: A latent variable framework for analysis and power estimation. Psychological Methods, 2, 371-402.

Saris, W. E., & Satorra, A. (1993). Power calculation in structural equation models. In K. A. Bollen & J. S. Long (Eds.), Testing structural equation models (pp. 181-204). Newbury Park, CA: Sage.

Satorra, A., & Saris, W. E. (1985). The accuracy of a procedure for calculating the power of the likelihood ratio test as used within the LISREL framework. In C. O. Middendorp (Ed.), Sociometric research 1982 (pp. 129-190). Amsterdam: Sociometric Research Foundation.

Shadish, W. R., Cook, T. D., & Campbell, D. T. (2002). Experimental and quasi-experimental designs for generalized causal inference. Boston, MA: Houghton Mifflin Company.

Stevens, J. P. (2002). Applied multivariate statistics for the social sciences (4th ed.). Mahwah, NJ: Lawrence Erlbaum Associates.

Ware, J. H. (1985). Linear models for the analysis of longitudinal studies. American Statististician, 39(2), 95-101.

Willet, J. B., & Sayer, A. G. (1996). Cross-domain analyses of change over time: Combining growth modeling and covariance structure analysis. In G. A. Marcoulides & R. E. Schumacker (Eds.), Advanced structural equation modeling: Issues and techniques (pp. 125-157). Mahwah, NJ: Lawrence Erlbaum Associates.

Table 1 Simulated Data Conditions for Two Groups

Correlation Pattern of Repeated Measurements	Y1	1.00				
	Y2	.66	1.00			
	Y3	.59	.66	1.00		
	Y4	.52	.59	.66	1.00	
	Y5	.45	.52	.59	.66	1.00
Common STDs for Two Groups		10.00	10.00	10.00	10.00	10.00
Means of Group 1:		50.00	52.50	55.00	57.50	60.00
Means of Group 2:						
I Intercept Difference						
Small ES (d = .2)		52.00	54.00	57.00	59.00	62.00
Medium ES (d = .5)		55.00	57.50	60.00	62.50	65.00
Large ES (d = .8)		58.00	60.50	63.00	65.50	68.00
II Slope Difference						
Small ES (d = .2)		50.00	53.00	56.00	59.00	62.00
Medium ES (d = .5)		50.00	53.75	57.50	61.25	65.00
Large ES (d = .8)		50.00	54.50	59.00	63.50	68.00
III Intercept (d = .2) and Slope Differences (Ordinal Pattern-1)						
End with:						
Medium ES (d = .5)		52.00	55.25	58.50	61.75	65.00
Large ES (d = .8)		52.00	56.00	60.00	64.00	68.00
IV Intercept and Slope Differences (Ordinal Pattern-2)						
Start with:						
Medium ES (d = .5)		45.00	48.25	51.50	54.75	58.00
Large ES (d = .8)		42.00	46.00	50.00	54.00	58.00
V Intercept and Slope Differences (Disordinal Pattern)						
End with:						
Medium ES (d = .2)		52.00	53.50	55.00	56.50	58.00
Large ES (d = .5)		52.00	52.75	53.50	54.25	55.00

Table 2 Non-Converging Samples in SEM Latent Growth Modeling (%)

	Models				
Sample N	I	II	III	IV	V
50	1.60	0.58	0.60	0.63	8.33
100	0.32	0.03	0.05	0.00	6.55
200	0.03	0.00	0.00	0.00	3.25
300	0.00	0.00	0.00	0.00	2.30
400	0.00	0.00	0.00	0.00	1.93
500	0.00	0.00	0.00	0.00	0.98
600	0.00	0.00	0.00	0.00	1.03
700	0.00	0.00	0.00	0.00	0.60
800	0.00	0.00	0.00	0.00	0.68
900	0.00	0.00	0.00	0.00	0.43
1000	0.00	0.00	0.00	0.00	0.30

Figure Captions:

Figure 1 Unconditional (A) and Conditional (B) Latent Growth Models

Figure 2 Simulated Five Models of Group Differences

Model I: Group Difference in Intercept Only

Model II: Group Difference in Slope Only

Model III: Group Difference in both Intercept and Slope - Ordinal Group*Time Interaction Pattern (Start with Small ES; End with Medium or Large ES)

Model IV: Group Difference in both Intercept and Slope - Ordinal Group*Time Interaction Pattern (Start with Medium or Large ES; End with Small ES)

Model V: Group Difference in both Intercept and Slope - Disordinal Group*Time Interaction Pattern (Start with Small ES; End with Medium or Large ES in Reversed Direction)

Figure 3 Model I: Type I Error for Falsely Declaring Group Intercept and Slope Differences

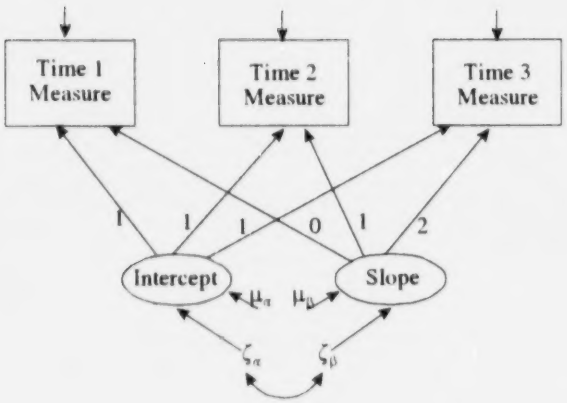
Figure 4 Model I: Power for Detecting Intercept Difference: SEM and Univariate Repeated ANOVA (Assuming No Slope Difference)

Figure 5 Model II: Power for Detecting Slope Difference (Zero Intercept Difference; End with $\underline{d} = 0.2$, $\underline{d} = 0.5$, or $\underline{d} = 0.8$)

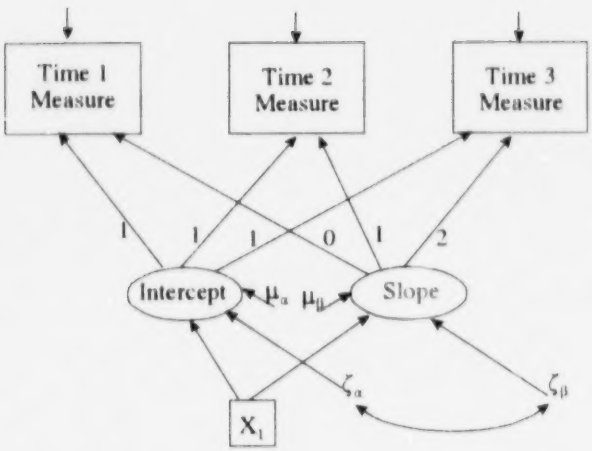
Figure 6 Model III (Ordinal Interaction Pattern 1): Power for Detecting (a) Intercept Difference ($\underline{d} = 0.2$), and (b) Slope Differences: End with $\underline{d} = 0.5$, or $\underline{d} = 0.8$

Figure 7 Model IV (Ordinal Interaction Pattern 2): Power for Detecting Slope Differences: Start with $d = 0.5$, or $d = 0.8$; End with $\underline{d} = 0.2$

Figure 8 Model V (Disordinal Interaction Pattern): Power for Detecting Slope Differences – Start with $\underline{d} = 0.2$; End with $\underline{d} = 0.2$, or $\underline{d} = 0.5$ in Reversed Direction



Model A



Model B

Figure 1

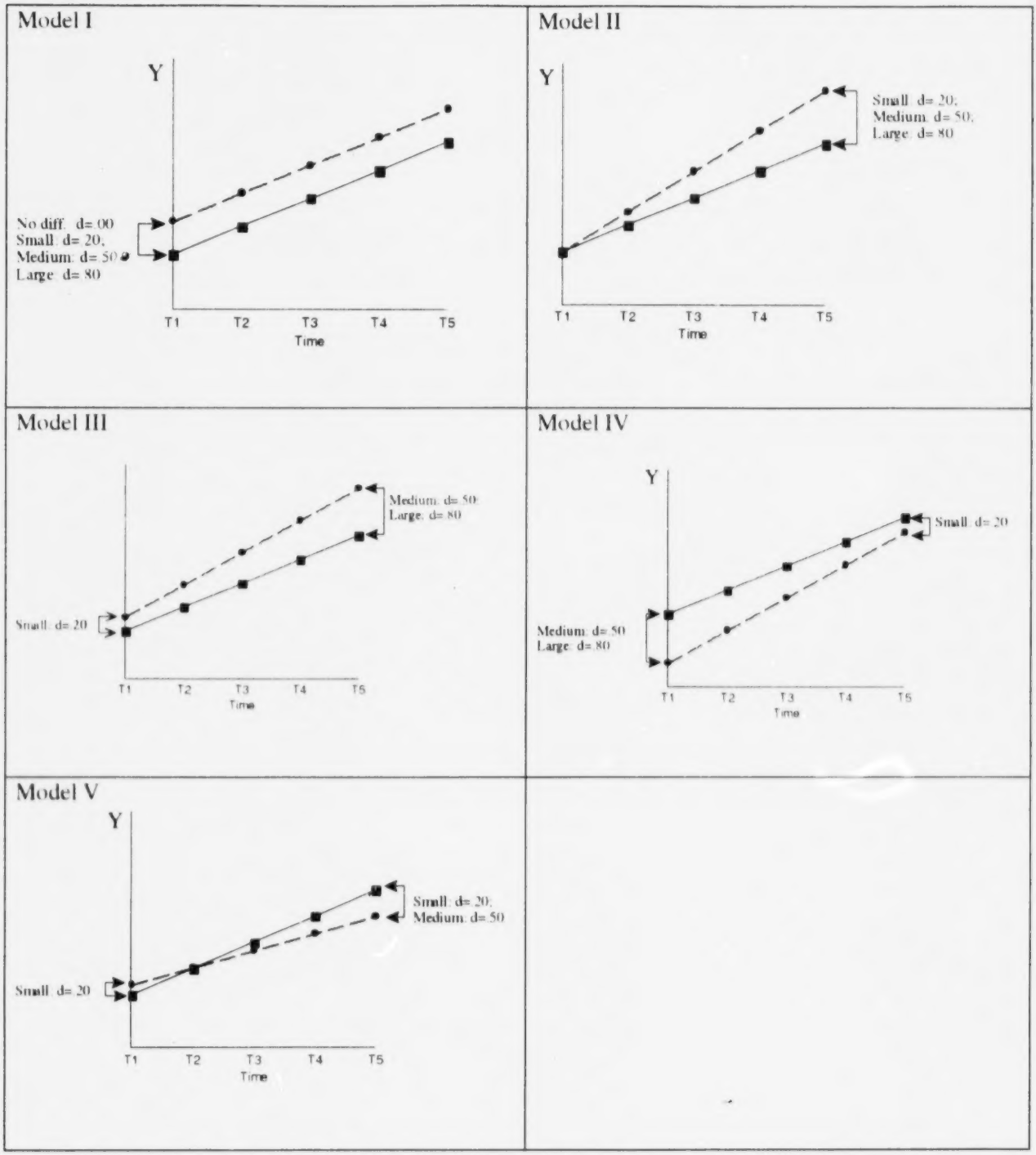

Model I
Y
No diff: d= 00
Small: d= 20;
Medium: d= 50
Large: d= 80
T1
T2
T3
T4
T5
Time
Model II
Y
Small: d= 20;
Medium: d= 50;
Large: d= 80
T1
T2
T3
T4
T5
Time
Model III
Medium: d= 50
Large: d= 80
Small: d= 20
T1
T2
T3
T4
T5
Time
Model IV
Y
Small: d= 20
Medium: d= 50
Large: d= 80
T1
T2
T3
T4
T5
Time
Model V
Y
Small: d= 20;
Medium: d= 50
Small: d= 20
T1
T2
T3
T4
T5
Time

Figure 2

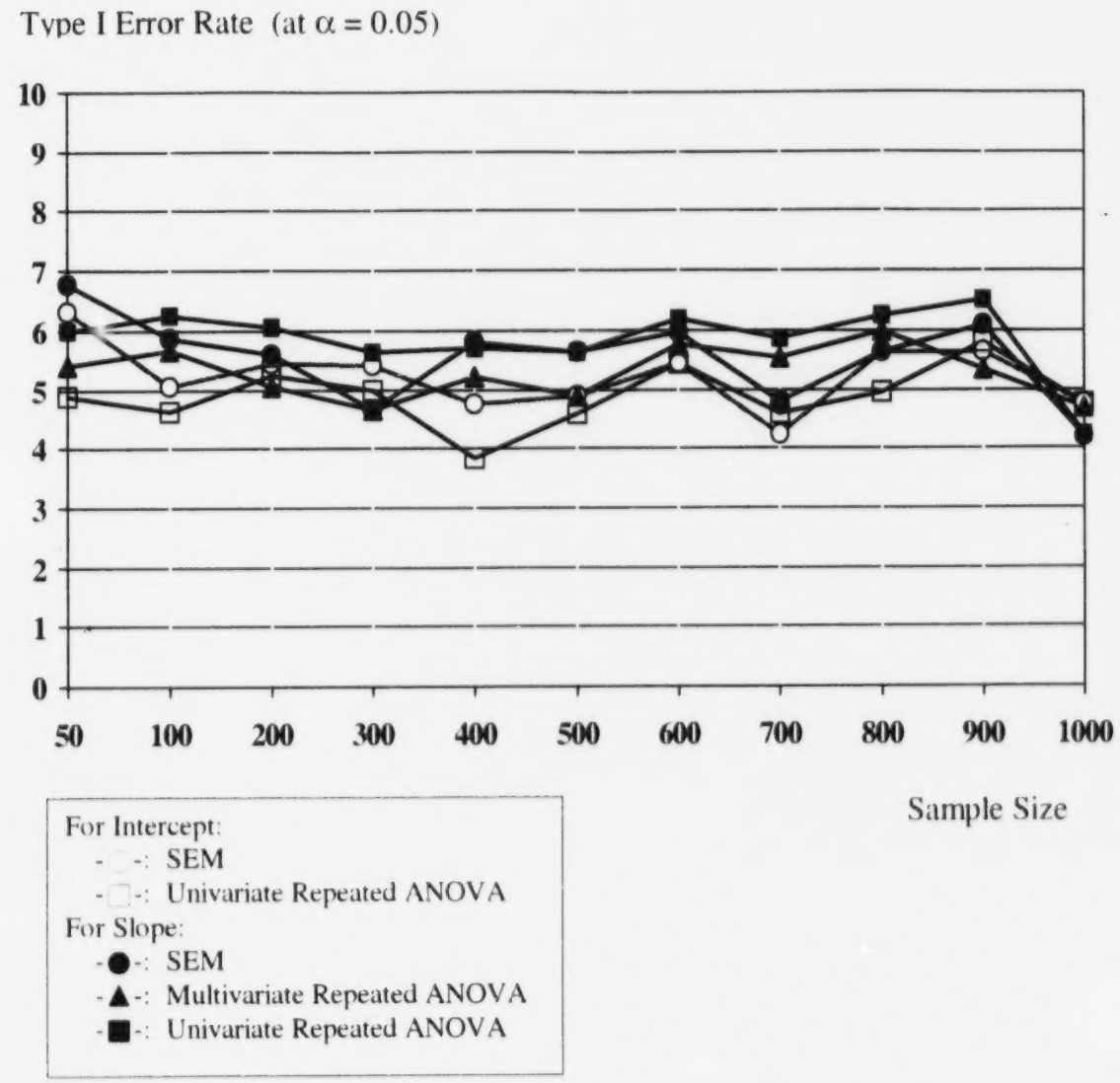

Type I Error Rate (at α = 0.05)
10
9
8
7
6
5
4
3
2
1
0
50
100
200
300
400
500
600
700
800
900
1000
Sample Size
For Intercept:
-□-: SEM
-□-: Univariate Repeated ANOVA
For Slope:
-●-: SEM
-▲-: Multivariate Repeated ANOVA
-■-: Univariate Repeated ANOVA

Figure 3

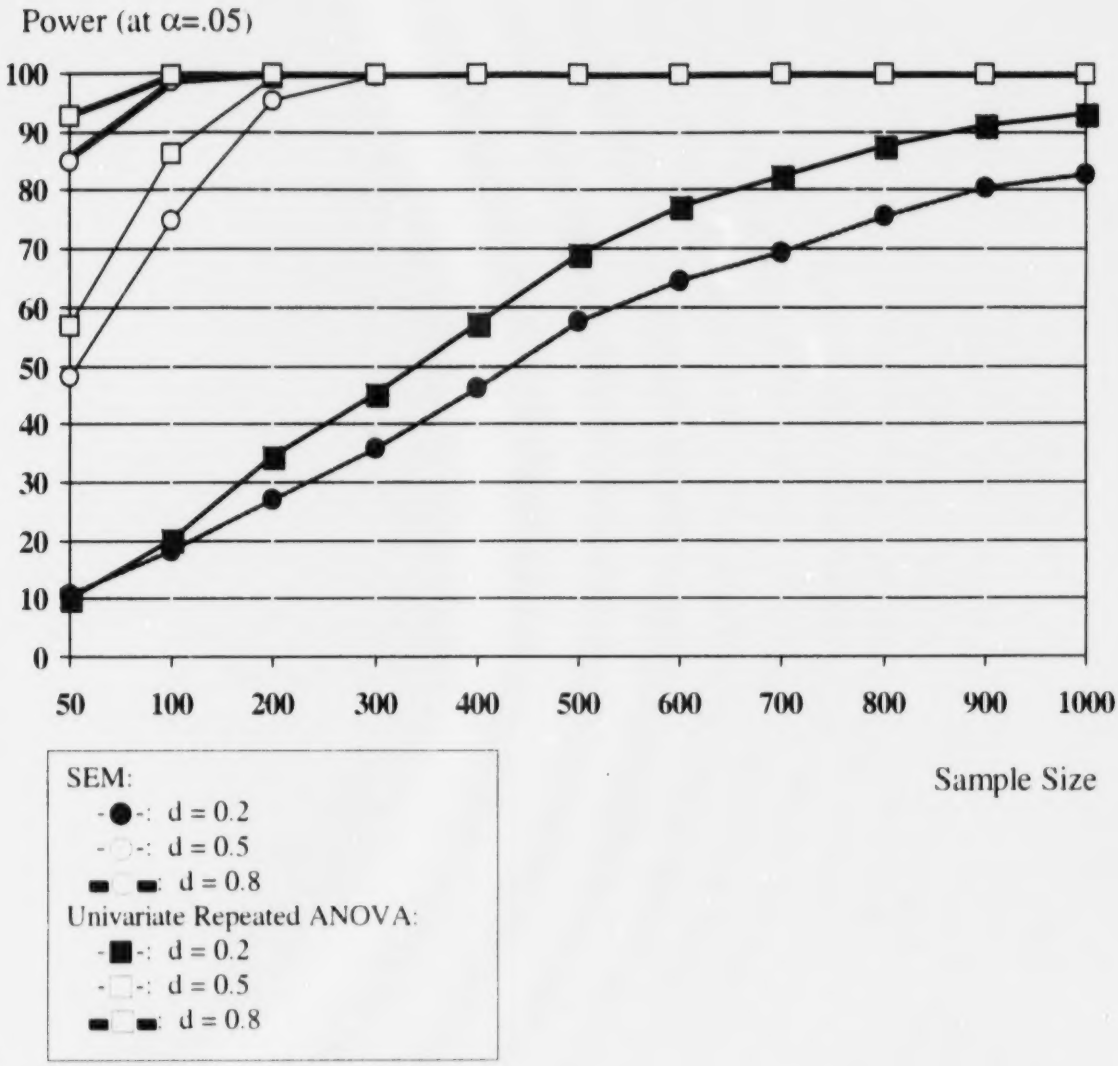

Power (at α=.05)
100
90
80
70
60
50
40
30
20
10
0
50
100
200
300
400
500
600
700
800
900
1000
Sample Size
SEM:
d = 0.2
d = 0.5
d = 0.8
Univariate Repeated ANOVA:
d = 0.2
d = 0.5
d = 0.8

Figure 4

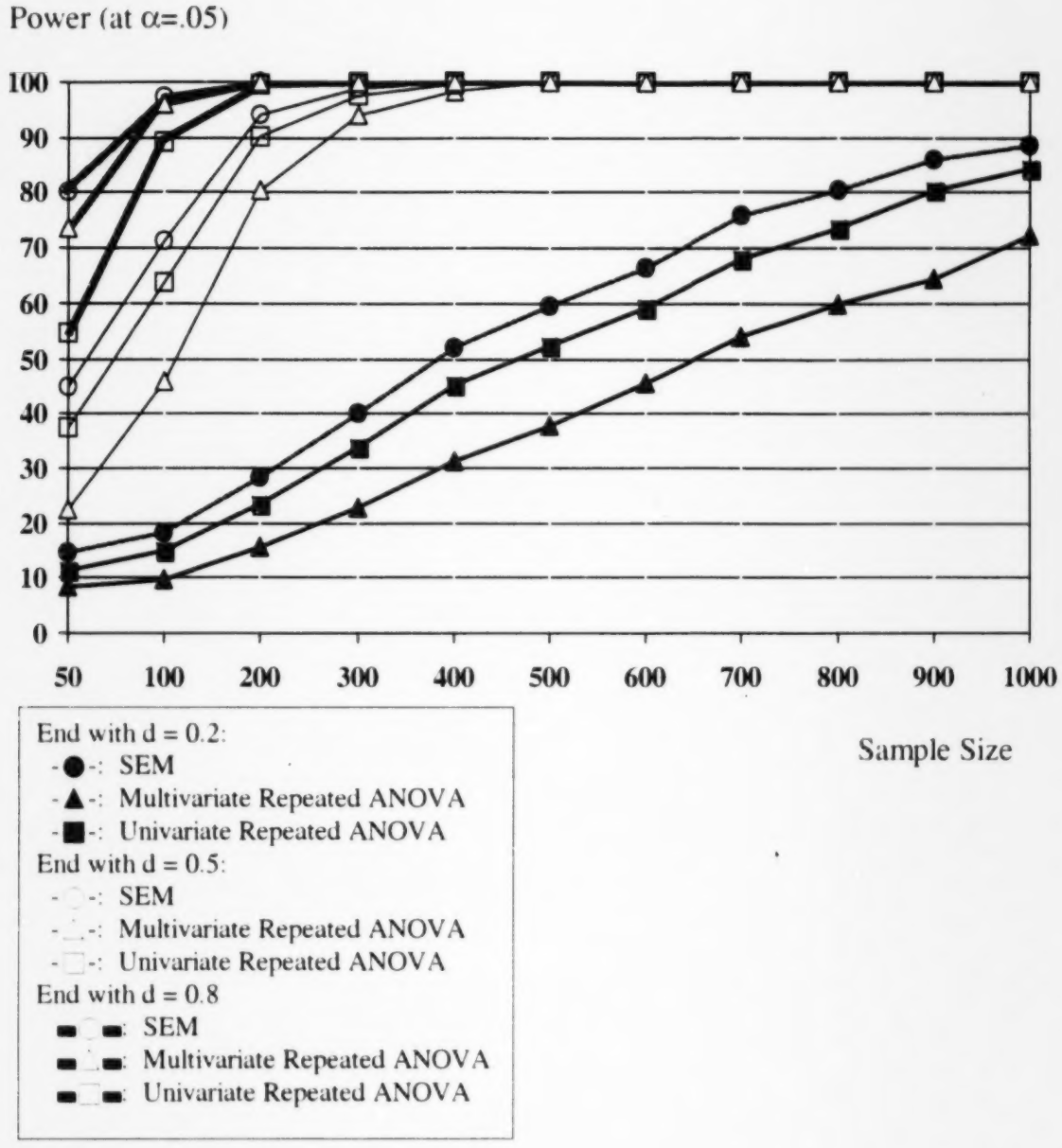

Power (at α=.05)
100
90
80
70
60
50
40
30
20
10
0
50
100
200
300
400
500
600
700
800
900
1000
Sample Size
End with d = 0.2:
-●-: SEM
-▲-: Multivariate Repeated ANOVA
-■-: Univariate Repeated ANOVA
End with d = 0.5:
-○-: SEM
-△-: Multivariate Repeated ANOVA
-□-: Univariate Repeated ANOVA
End with d = 0.8
SEM
Multivariate Repeated ANOVA
Univariate Repeated ANOVA

Figure 5

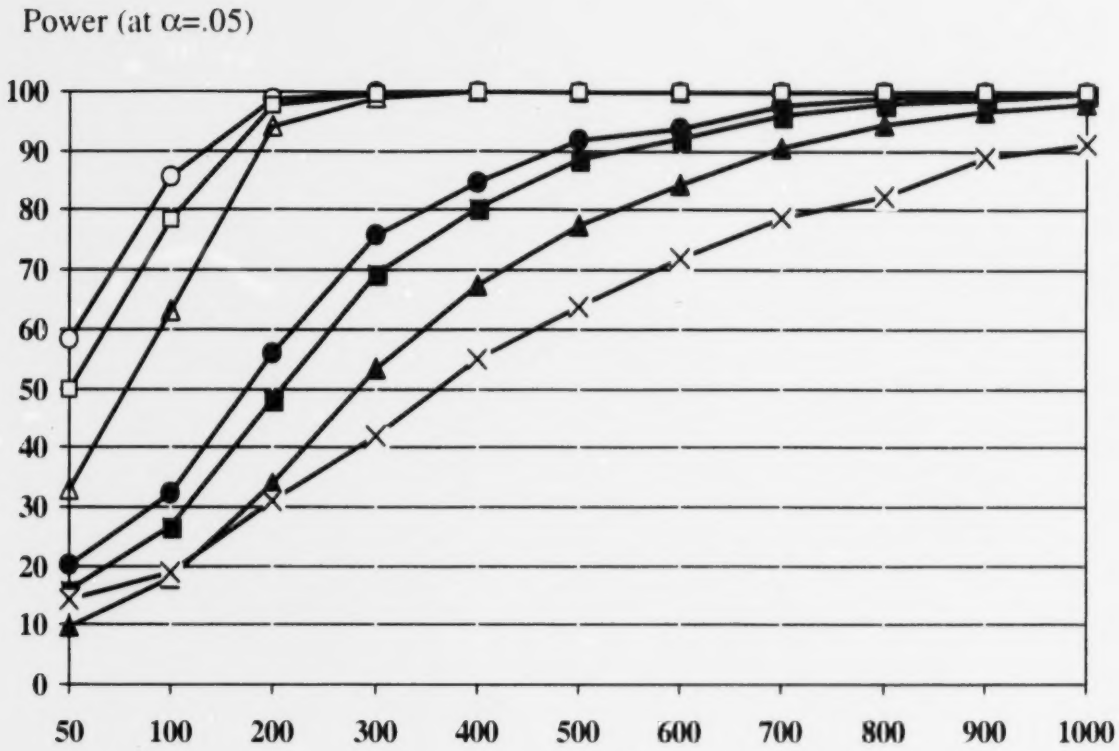

Power (at α=.05)
100
90
80
70
60
50
40
30
20
10
0
50
100
200
300
400
500
600
700
800
900
1000
Sample Size
For Slope Difference:
(1) End with d = 0.50:
-●-: SEM
-▲-: Multivariate Repeated ANOVA
-■-: Univariate Repeated ANOVA
(2) End with d = 0.80:
-□-: SEM
-△-: Multivariate Repeated ANOVA
-□-: Univariate Repeated ANOVA
For Intercept Difference:
- · -: SEM

Figure 6

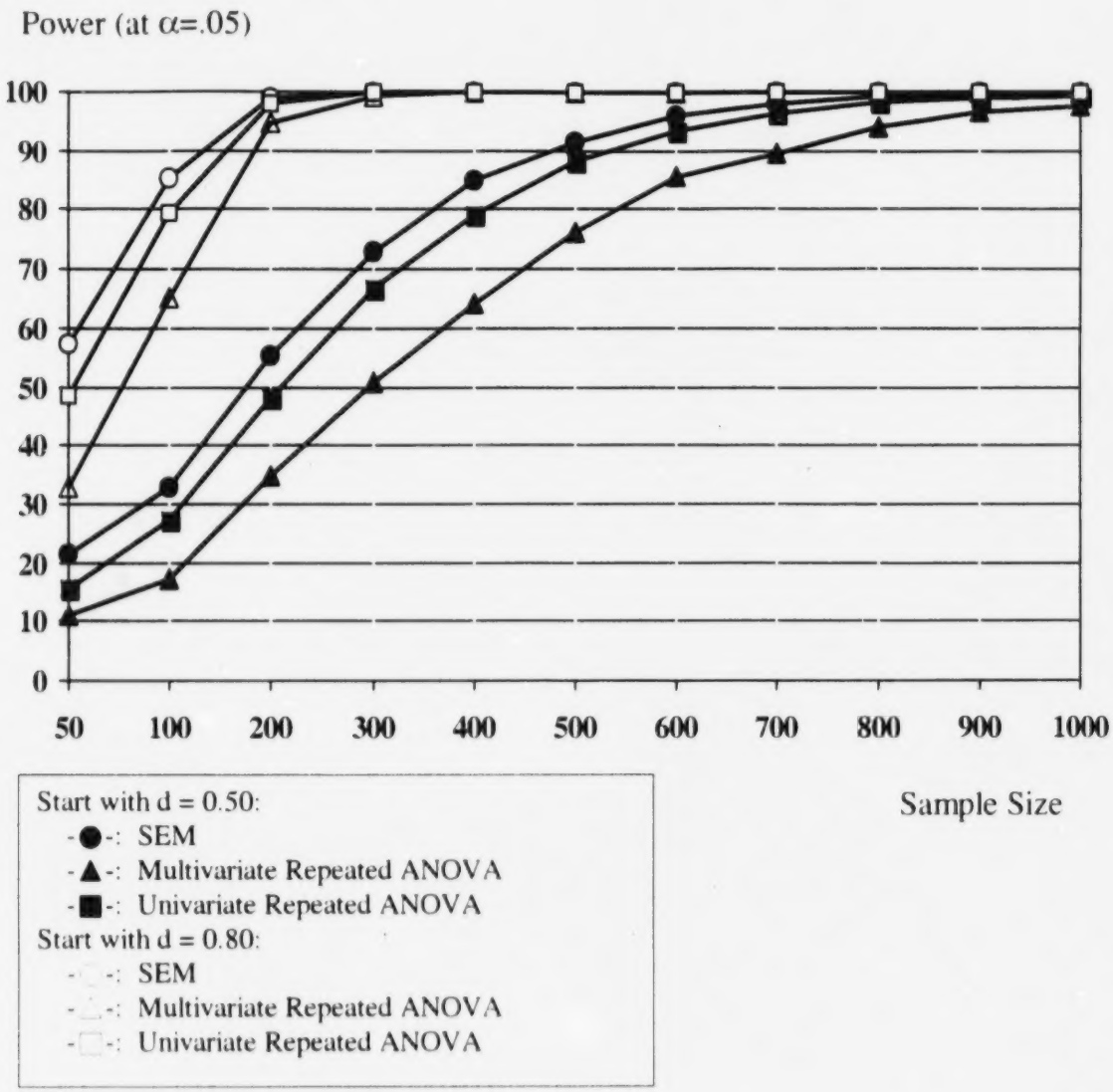

Power (at α=.05)
100
90
80
70
60
50
40
30
20
10
0
50
100
200
300
400
500
600
700
800
900
1000
Sample Size
Start with d = 0.50:
-●-: SEM
-▲-: Multivariate Repeated ANOVA
-■-: Univariate Repeated ANOVA
Start with d = 0.80:
-○-: SEM
-△-: Multivariate Repeated ANOVA
-□-: Univariate Repeated ANOVA

Figure 7

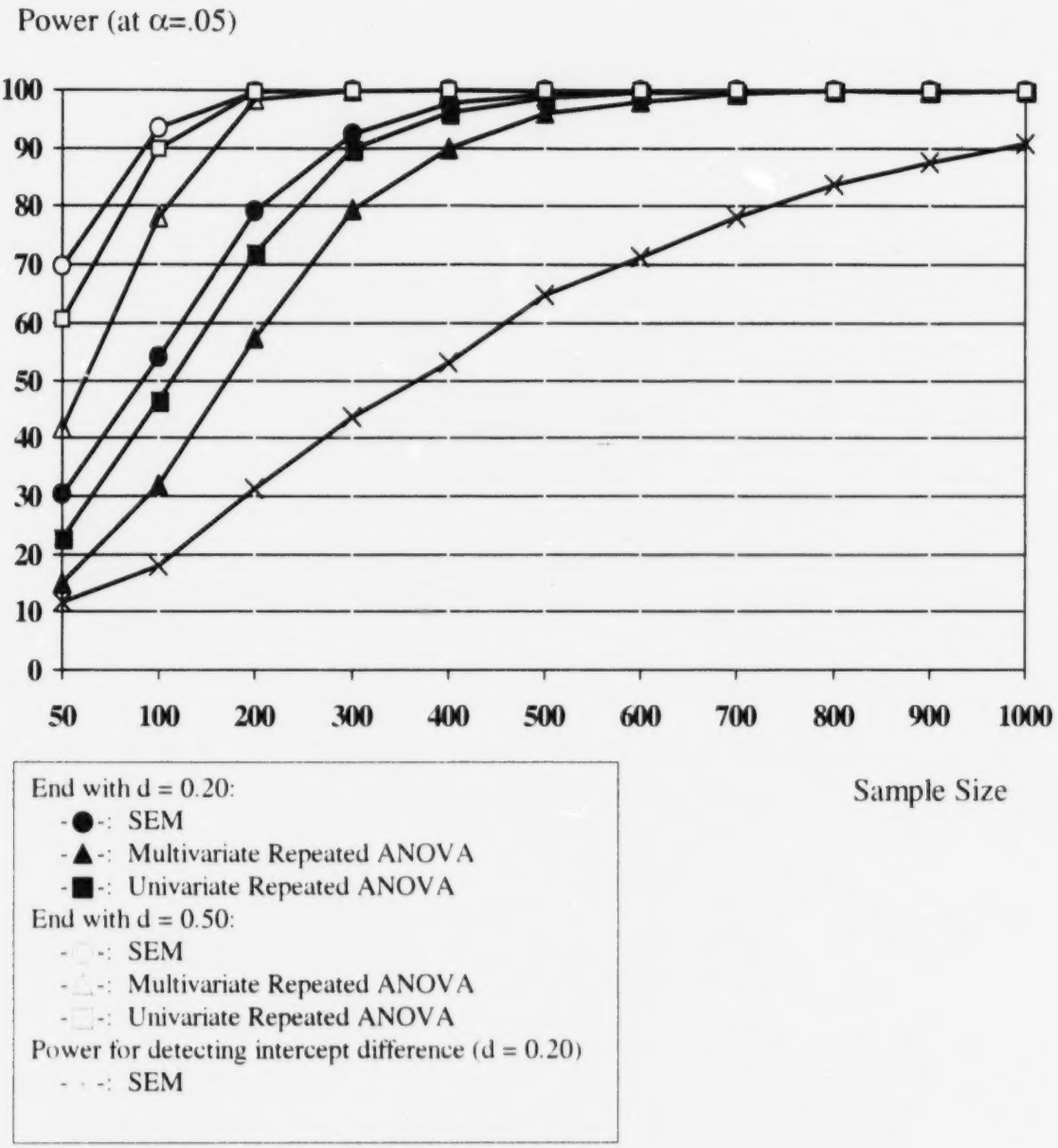

Power (at α=.05)
100
90
80
70
60
50
40
30
20
10
0
50
100
200
300
400
500
600
700
800
900
1000
Sample Size
End with d = 0.20:
-●-: SEM
-▲-: Multivariate Repeated ANOVA
-■-: Univariate Repeated ANOVA
End with d = 0.50:
-○-: SEM
-△-: Multivariate Repeated ANOVA
-□-: Univariate Repeated ANOVA
Power for detecting intercept difference (d = 0.20)
- - -: SEM

Figure 8